# UC Berkeley
## Proceedings of the Annual Meeting of the Berkeley Linguistics Society

**Author**
Esling, John H

The Articulatory Function of the Larynx and the Origins of Speech
Author(s): John H. Esling
*Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society* (2014), pp. 121-150
General Session and Thematic Session on Language Contact
Editors: Kayla Carpenter, Oana David, Florian Lionnet, Christine Sheil, Tammy Stark, Vivian Wauters

# The Articulatory Function of the Larynx and the Origins of Speech

JOHN H. ESLING
*University of Victoria*

## Introduction

The 'laryngeal articulator,' consisting of the glottal mechanism, the supraglottic tube, the pharyngeal/epiglottal mechanism, and including three levels of folds: the vocal folds, the ventricular folds, and the aryepiglottic folds, is shown to be responsible for the generation of multiple source vibrations and for the complex modification of the pharyngeal resonating chamber that accounts for a wide range of contrastive auditory qualities. These qualities are observed in a surprisingly large number of the languages of the world, both linguistically and paralinguistically, and they account for sounds which have been labeled as 'pharyngeal,' as 'epiglottal,' and as various phonation types. They reflect an expanding range of what have been known as the 'states of the glottis' and which may be more properly termed 'states of the larynx.'

It has also been observed that infants, in their first months of life, produce a range of qualities, reflecting both phonatory possibilities and stricture types, that can also be attributed to the laryngeal articulator mechanism. Systematic observation of infants' early speech production reveals that the control of articulatory detail in the pharynx is mastered during the first year of life. Understanding and control of manner of articulation in the pharynx appears to be a prerequisite for expanding articulatory control into the oral vocal tract. Taking the pharynx as a starting point for the ontogenetic learning of the speech production capacity offers fruitful insights into the phylogenetic development of speech.

Our research into the earliest vocalizations by infants in a research project including English (Victoria), French (Paris), Arabic (Morocco), and Bai (China) shows that (1) speech begins in the pharynx, (2) the production of phonation begins with laryngeally constricted settings, (3) infants explore their phonetic production capacity by employing 'dynamic alternations' and 'pharyngeal

priming'. Evidence from the Infant Speech Acquisition (InSpA) Project illustrates instances of 'phonetic play' that demonstrate how infants systematically acquire basic control over the speech mechanism and over the arrays of place and manner of articulation during their first year of life.

## 1 Laryngeal Articulation

The 'Laryngeal Articulator,' consisting of the glottal mechanism, the supraglottic tube, the pharyngeal/epiglottal mechanism, and including three levels of folds: the vocal folds, the ventricular folds, and the aryepiglottic folds, is shown to be responsible for the generation of multiple source vibrations and for the complex modification of the pharyngeal resonating chamber that accounts for a wide range of contrastive auditory qualities. In a program of experimental phonetic research spanning two decades, it has been shown that the laryngeal articulator plays a key role in differentiating linguistic phonetic meaning in a surprisingly large number of languages of the world (Esling 1996, Esling and Edmondson 2002, Esling, Fraser, and Harris 2005, Edmondson, Padayodi, Hassan and Esling 2007). In the Laryngeal Articulator Model (Esling 1999, 2005, Edmondson and Esling 2006, Catford and Esling 2006), it has been demonstrated that the glottis is not the only source of periodic energy in the larynx, that the aryepiglottic folds also generate vibrations attested in speech sounds, and that the laryngeal constrictor, as the shaper of pharyngeal articulations and of pharynx volume, contributes primary cavity resonance in 'tonal register' and 'vowel harmony' sound systems. Many languages of the world exhibit features that can be classified in terms of 'laryngeal quality'. The acoustic cues of these features illustrate an extensive range of use of the pharyngeal resonator and the laryngeal constrictor mechanism (controlling changes from the glottis through the aryepiglottic folds). Elements of the fine control of laryngeal constriction have been observed laryngoscopically in over 20 languages and modeled to illustrate the parameters of movement available in the laryngeal/pharyngeal space. Laryngoscopic evidence drawn from e.g. Tibeto-Burman, Semitic, Cushitic, Kwa, and Gur languages demonstrates the distinctive use of the laryngeal articulator in pharyngeal trilling combined with glottal voicing, voiceless pharyngeal trilling, and epilaryngeal tube shaping to create opposing vocal register series. One such series is the [ATR/-ATR] contrast, but other harmony systems share a similar phonetic basis.

The laryngeal articulator has also been identified as the principal articulator that infants first start to control as they test and practice their phonetic production skills from birth through the first several months of life. The auditory/acoustic cues generated in the pharynx in the range of languages we have observed are the same elements of sound production observed in early infancy. The infant vocalization data illustrate that laryngeal quality is primal, that control of the articulatory and perhaps acoustic cues of speech originates in the pharynx, and

that the acquisition of the ability to produce manners of articulation spreads from the pharynx in a process that parallels and complements the ability of infants to discriminate speech sound categories perceptually. It is important to point out that our initial observations relating the sounds of infant speech to the pharynx were auditory observations based on working with the various adult phonetic realizations across the languages of the world and then listening to infant productions that very clearly resembled a large number of the adults' pharyngeal forms in great detail. The audio samples presented in the BLS paper are designed to illustrate this close relationship in auditory quality.

## 2       The Laryngeal Articulator Model

### 2.1       The Two-Part Vocal Tract

In the laryngeal articulator view, pharyngeals are considered to be laryngeal articulations, i.e. a function of the aryepiglottic sphincter mechanism, rather than being primarily a function of tongue position. The tools of investigation that we use in our phonetic research have allowed us to formulate a theory whereby the production of pharyngeal sounds is isolated within the 'laryngeal articulator'. We have examined laryngeal articulation in over 20 languages from diverse language families across the world to demonstrate the laryngeal states (Esling 2006), phonatory postures (Catford and Esling 2006), and manners of articulation (Carlson, Esling, and Harris 2004, Edmondson, Esling, Harris, and Huang 2005) that the laryngeal articulator can produce. We have shown that various complex adjustments within a relatively simple mechanism – the laryngeal constrictor – are responsible for the production of multiple manners of pharyngeal consonants, secondary sound source vibrations, and changes in pharyngeal cavity resonance, and we have shown how some languages use these articulatory parameters to generate distinctions that have been characterized phonologically as [lax/tense] or [+ATR/-ATR] (Edmondson, Padayodi, Hassan, and Esling 2007).

The vocal tract is reconceptualized into two primary articulatory domains in the Laryngeal Articulator Model (Esling 2005). The laryngeal section of the vocal tract is conceived of as a series of valves, each of which is responsible for a range of articulatory configurations and is made up of its own unique anatomical structures (Edmondson and Esling 2006). Figure 1 illustrates this conceptual expansion of the articulatory capacity of the laryngeal vocal tract and the primacy of the aryepiglottic folds in the production of sounds that have been termed pharyngeals and epiglottals. For infant speech, regardless of the differences in vocal tract shape from adults, it will be important to identify accurately sounds that originate in the laryngeal sector.

**Figure 1:** The 'two-vocal-tract' reconceptualization of the standard articulatory model of the vocal tract (minus the nasal tract) with jaw height determining close/open vowel setting at the front, and tongue raising/retracting determining vowel quality at the back. T, tongue; U, uvula; E, epiglottis; H, hyoid bone; AE, aryepiglottic folds; A, arytenoid cartilages; VF, vocal folds; Th, thyroid cartilage; C, cricoid cartilage. This figure is reproduced from Esling's original 2005 work in the *Canadian Journal of Linguistics*.
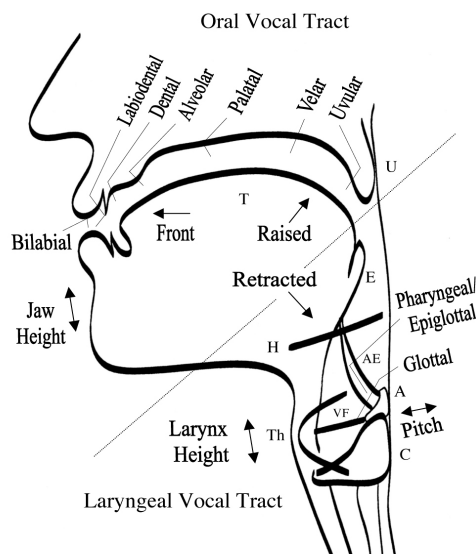


## 2.2 Laryngoscopic Descriptions of Laryngeal Behavior

Using this model, we have determined, for example, what the mechanism of pharyngeal trilling involves and where it occurs in the speech sounds of a language with pharyngeals in its phonology. We have also identified the mechanism of pharyngeal shaping that generates contrasting laryngeal features such as [+ATR/-ATR] vowels. The supplementary source mechanism and the resonance features resulting from narrowing of the pharyngeal space are illustrated particularly well in studies of Bai, Iraqi Arabic, Somali, Yi, Akan, and Kabiye (Edmondson, Padayodi, Hassan, and Esling 2007). Laryngeal/pharyngeal production in each language has been studied articulatorily by means of transnasal laryngoscopy. Bai was chosen because of its vocal register contrast that involves a complex of glottal and supraglottic phonatory modification (Esling and Edmondson 2002); Iraqi Arabic because of its extreme and phonetically challenging pharyngeal reflexes (Heselwood 2007, Hassan, Esling, Moisik, and Crevier-Buchman 2011); Somali because it has pharyngeals as well as a vocalic harmony system that interacts with shaping in the pharynx; Yi because of its

register series that do not induce phonatory contrasts (Edmondson, Esling, Lama, Harris, and Li 2001); Akan because much is known about its phonetics, and because it is historically the model language for 'ATR' harmony (Tiede 1996). Kabiye, a Gurunsi language of Togo, Ghana, and Benin, compares closely with Akan with extremely regular vowel series and virtually exceptionless adherence to the rules of 'ATR' vowel harmony.

The experimental phonetic equipment in our laryngoscopic research facility consists of a Kay Elemetrics Rhino-laryngeal-stroboscope (RLS 9100) with a constant halogen cold light source to photograph the actions of the larynx. An Olympus ENF-P3 fibreoptic nasendoscope is attached to the camera (Panasonic GP-US522) and light source with a 28mm lens for optimal wide-angle framing of larynx mechanisms during extreme pharyngeal articulations and of laryngeal postures during the varying pitch conditions in which tonal paradigms occur. Recordings, originally made on VHS tapes, are now made directly on a Sony DCRTRV17 Mini-DV Digital Camcorder with high-quality audio. Video images have been postprocessed using Adobe Premiere 6.5 software and, more recently, with Sony Vegas Pro.

Canonical phonetic profiles obtained in baseline research (Esling 1996, 1999, 2005, 2006, Esling and Harris 2005) serve as the basis for comparing the production of phonemic contrasts by native-speaker subjects in the array of languages we have studied. Languages in the video database of the larynx and pharynx include: Nuuchahnulth (Wakashan), Nlaka'pamuxcín (Salish), Tigrinya (Semitic), Palestinian Arabic (Semitic), Iraqi Arabic (Semitic), Somali (Cushitic), Amis (Austronesian), Yi (Tibeto-Burman), Bai (Sino-Tibetan/Tibeto-Burman), Tibetan (Tibeto-Burman), Sui (Kam-Daic), Thai (Daic), Pame (Oto-Manguean), Cantonese (Sinitic), Chinese (Sinitic/whisper studies), Danish (Germanic), English (Germanic/whisper studies), Korean (Altaic), Bor Dinka (Nilotic), Chong (Mon-Khmer), Akan (Niger-Congo, Kwa), and Kabiye (Niger-Congo, Gur).

## 2.3    Laryngeal Qualities and Place and Manners

In the same way that canonical phonetic profiles obtained in baseline research constitute the basis for comparing sounds produced by native-speaker subjects of various languages, the data that we have obtained from the various native-speaker subjects constitute the basis for comparing sounds produced by infants in the Infant Speech Acquisition (InSpA) Project. Whereas invasive laryngoscopic experimentation can be carried out with adult subjects, invasive procedures are not possible with infant subjects; and they are not needed. The principal phonetic classification that takes place in the adult language studies remains largely auditory, with instrumental data providing support for the auditory distinctions that we identify in the various phonologies. The phonetic classification that takes place in the infant language studies is wholly auditory. Classifications are

developed and assigned on the basis of intensive training in the system of the Laryngeal Articulator Model, with which all members of the research team are intimately familiar. Categories follow those developed while working with the data from the array of languages listed above and which are taught in undergraduate classes and in postgraduate seminars at the University of Victoria.

**Figure 2:** Maximum stricture of medial epiglottal stop [ʡː] in Iraqi Arabic /faʕʕal/ 'made active' (left) and voiceless epiglottal trill [ʜː] in /saħħar/ 'made magic' (right). Aryepiglottic folds, middle.
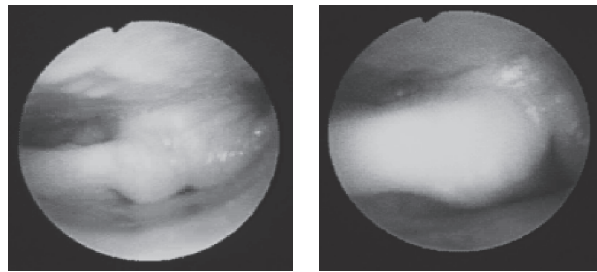


The categories of analysis from languages observed laryngoscopically include qualities and place and manner designations. Some examples of categories from our previous research are epiglottal stop (widely distributed in Semitic, Cushitic, Wakashan, Salish, and Austronesian languages), glottal stop (with an even wider distribution), voiceless pharyngeal trill (in Iraqi Arabic and Somali as well as in Tibeto-Burman register). In Somali, in a voiceless pharyngeal, an effect of harsh trilling of the aryepiglottic folds co-occurs, and it also appears that the combination of tight, harsh phonation at the glottis propagates through the supraglottic tube, inducing vibrations of the narrowed structures above it, including the tip of the epiglottis. A pharyngeal tap has also been identified in our data from Iraqi Arabic (a speech sound category not yet attested in the chart of the IPA). Most of these sound types, in fact, had been underattested among the world's languages prior to the results of our laryngoscopic research; so it is now possible to say that many if not most of the categories of sounds generated by the laryngeal constrictor occur far more commonly than had been previously thought. It is also possible to specify more clearly how each category is related articulatorily to the others. Examples of the posture of an epiglottal stop and of a point during voiceless aryepiglottic trilling are shown in Figure 2. In both, the aryepiglottic folds are raised up and forwards towards the retracting tongue and epiglottis. During the trill, the aryepiglottic folds are vibrating vigorously while the glottis is open beneath in the state of breath. Detailed studies of this vibratory phenomenon as a secondary laryngeal periodic source have appeared in Moisik, Esling, and Crevier-Buchman (2010) and in Hassan, Esling, Moisik, and Crevier-

Buchman (2011). An articulatory model has also been developed which compares the movement at the glottis with the movement at the aryepiglottic folds (Moisik 2008).

The qualities (due to pharyngeal resonance) that we have identified as a function of the laryngeal articulator are an example of how source generation and articulatory dynamics in the laryngeal space interact with other phonological specifiers that are not considered to be 'pharyngeal' per se. The 15-way tonal register syllable distinction in Bai (Tibeto-Burman) based on pitch, phonation type, laryngeal-constrictor tension, and nasal quality is a good example of this relationship (Edmondson, Esling, Lama, Harris, and Li 2001, Esling and Edmondson 2002). The Bai registers have been called 'lax' and 'tense,' which are good phonological terms to differentiate the qualities. Phonetically speaking, we have argued that 'lax' reflects [-cet] ([-constricted epilaryngeal tube]) syllabic register, while 'tense' reflects [+cet] ([+constricted epilaryngeal tube]) syllabic register (Moisik and Esling 2011b). The feature [+cet] represents the engagement of the laryngeal constrictor mechanism and results in qualities that are articulatorily the same as pharyngealization. Yi (Tibeto-Burman) is an example of a language whose vocalic system is characterized primarily by a [-cet/+cet] syllabic register contrast, which is also called lax/tense. In phonetic terms, tense recapitulates the engagement of the laryngeal constrictor mechanism such that the resulting resonance of the pharyngeal tubes generates what has been called 'raised larynx voice' (Laver 1980). Lax encompasses an opposite posture – neutral larynx height with an open epilaryngeal tube, or even lowered larynx. But this is not an areally isolated phenomenon. West African languages that have been labeled with [+ATR/-ATR] harmony series demonstrate the same contrast in laryngeal behavior as those Tibeto-Burman languages with [-cet/+cet] (lax/tense) distinctions. Dinka as well as Somali also share the use of this phonetic property but in a paradigm where quality plays a more subtle role. The West African 'ATR' paradigm is virtually identical in articulatory phonetic terms to the Tibeto-Burman lax/tense paradigm. Examples are Akan/Twi (Kwa) and Kabiye (Gur), which can each be thought of as having one series of vowels that is [-cet] (+ATR) [ i, e, u, o, a ], and a parallel set of vowels that is [+cet] (-ATR) [ ɪ, ɛ, ʊ, ɔ, ɑ ]. That the distinction is pharyngeal, i.e. a function of the laryngeal constrictor mechanism, was hypothesized in 1996 (Esling 1996) and eventually confirmed through laryngoscopic observation (Edmondson and Esling 2006, Edmondson, Padayodi, Hassan, and Esling 2007). The constricted series of these paradigms will be used here to make the connection between the possibilities that are available for contrastive purposes in adult phonologies and with the way that infants begin their articulatory phonetic experience. The laryngoscopic examples that follow are of adults, but the constricted postures are very similar to the innate shape of the infant vocal tract (Fitch and Giedd 1999).

In the [+cet] (-ATR) vowels in Akan, shown on the right in Figure 3, the flatter forward-bending aryepiglottic angle, the more retracted tongue and raised larynx, and the narrowed posture of the laryngeal constrictor reflect the constricted state of the laryngeal vocal tract. Anatomically speaking, the constricted posture on the right resembles somewhat the shape of the naturally straight and short infant vocal tract. This comparison will be relevant when discussing what kinds of sounds infants produce in their first months of life. The existence of this body of auditorily catalogued laryngoscopic data will also assist us in understanding what inventories of laryngeally generated sounds infants produce at what stages of vocal development. The same situation exists in Kabiye (Gur), which has a similar [+ATR/-ATR] vowel harmony series. Each one of the vowels in the [+cet] (-ATR) set [ ɪ, ɛ, ʊ, ɔ, ɑ ] demonstrates a tightened configuration of the laryngeal articulator, as seen in the image on the right in Figure 4.

**Figure 3:** Mid-initial-vowel laryngeal posture of Akan /midiʼ/ 'I am eating' (left) and /mɪ̠dɪ̀/ 'I am called' (right). The lines trace the aryepiglottic fold angle: low and open on the left; raised and forwards-closing (towards bottom of photo) on the right.
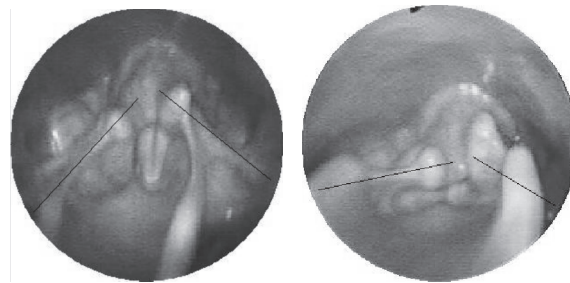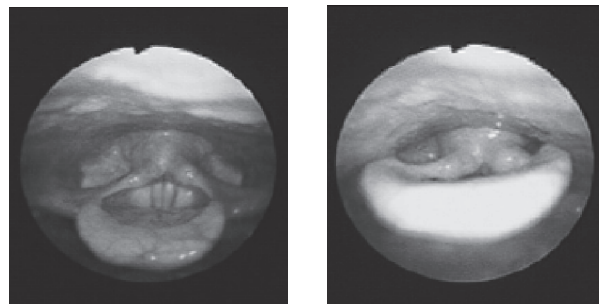


**Figure 4:** Mid-vowel laryngeal/pharyngeal posture of Kabiye /tú/ 'elephant' (left) and /tʊ́/ 'bee' (raised/constricted, right). Epiglottis, bottom; aryepiglottic folds, middle; pharyngeal wall, top.

## 2.4    Implications for the IPA Vowel System

In addition to consonantal place and manner, therefore, vowels are also affected by the action of the laryngeal constrictor. This is clear in the evidence from Akan and Kabiye, where vowel quality (noted by differing symbol shapes) changes as pharyngeal quality changes. The ostensibly redundant notation of adding a subscript retracting symbol captures the generalization of the pharyngeal quality common to this series of vowels (Padayodi 2008). The upwards and forward-flattening aryepiglottic angle, retracted tongue and raised larynx during laryngeal constriction and the consequent narrowed configuration of the pharyngeal resonator produce a vowel that is 'lower' and 'backer' on the vowel chart (i.e. retracted), whatever the individual vowel quality due to oral lingual position.

These considerations imply that the 2005 International Phonetic Alphabet vowel chart may need revision in terms of its traditional 'Front-Back' and 'Close-Open' dimensions. The issue revolves around identifying articulators. If the tongue is taken to be the only articulator defining vowel quality and the vowel space, then the effect introduced by the laryngeally constricted set of vowels in the West African phonologies reviewed here occupies an ill-defined place in the system. 'Back' does not account adequately for the effect of the laryngeal constrictor mechanism. Another parameter is needed in the deepest part of the vocal tract. And since opening of the jaw only applies to the 'Front' section of the vocal tract, 'Open' is not a particularly apt qualifier for what happens to vowels in the lower right region of the vowel chart. These implications for revision are discussed in detail in the new *Handbook of Phonetic Sciences* (Esling 2010).

In many familiar systems of vowel location, vowels are described as lingually high or low and front or back. This conceptualization implies a model of lingual movement within the dimensions of a square space – four-cornered in two-dimensional terms – with the tongue moving up or down and from front to back. The tongue is usually represented in this model as the articulator responsible for changes in vowel quality along the high-low and front-back dimensions. This can be called the H-L-F-B model. This image of the tongue moving high in the mouth or back in the mouth, however, does not conform with articulatory behavior or with the evidence we have discovered for West African languages or for Tibeto-Burman phonologies or even for the effects of pharyngeals on 'low back' vowels in Semitic languages, as seen for Iraqi Arabic or Tigrinya. Neither is the traditional H-L-F-B model as useful an image as it could be for understanding how sound quality is shaped by articulator movement, vocal tract postures, and resulting cavity resonances in a multi-faceted set of chambers such as the vocal tract.

The principal reason why the H-L-F-B model inadequately represents (and perhaps even mistakenly portrays) the phonetics of the vocal tract is that it assumes oral lingual articulator activity while virtually ignoring laryngeal articulator activity (not to mention ignoring the role of the jaw). To the extent that the H-L-F-B model is intended to account for auditory quality, it has also misinformed acoustic theory. The assumption that H-L-F-B movement of the tongue drives vowel quality is not entirely adequate in the light of what we have discovered about how the laryngeal articulator controls the pharyngeal resonator. In fact, the laryngeal articulator can be also be shown to relate indirectly to velo-pharyngeal and mandibular settings in addition to lingual movement. The key in the development of a revised paradigm is to integrate a laryngeal articulator component between the mechanism of glottal airflow and the oral/front vocal tract articulators.

**Figure 5:** A revised IPA vowel chart to reflect three-way tongue movement: fronting, raising, and retraction, where retraction is a function of the laryngeal constrictor mechanism (Esling 2005, Edmondson and Esling 2006, Esling 2010).
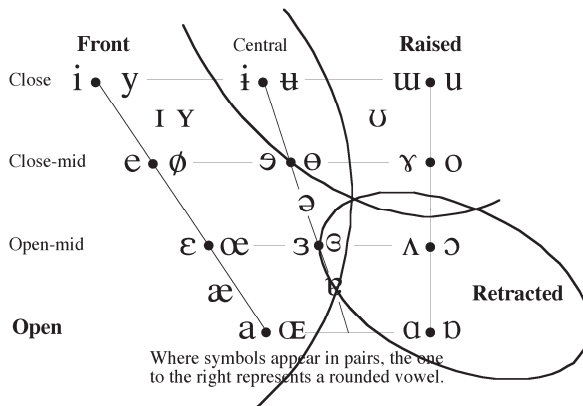


The development of the Laryngeal Articulator Model (Esling 2005) therefore prompts a revision of the conceptualization of the vowel chart, as suggested in Figure 5 (Esling 2010). At the front, the tongue may be 'Fronted,' but the articulator responsible for 'Open' quality is the jaw. At the back, it is the laryngeal constrictor. Tongue 'Raising' is also oral (and dorsal); but vowels affected by the action of the laryngeal constrictor will move in the direction of lingual 'Retraction,' which is the part of the laryngeal/pharyngeal constrictor mechanism that connects to the oral tract (cf. Figure 1). The principal action of laryngeal constriction (aryepiglottic fold sphinctering) precedes lingual retraction and larynx raising in the hierarchy of the complex maneuver (Edmondson and Esling 2006).

Larynx height is in fact a more voluntarily controllable variable than tongue retraction. Tongue retraction cannot occur without a prior engagement of the aryepiglottic sphincter mechanism. That is, there is an entailment relationship. The height of the larynx, however, can be controlled more freely. While reflexively linked to laryngeal constriction, so that the larynx normally raises during the constriction maneuver, the height of the larynx can be manipulated so that it may lower as a unit during constriction. Some phonologies may choose, as it were, to adopt this physiologically contrary setting as a coarticulatory component of pharyngeal sounds. The entailments at this vertical level of control, where aryepiglottic sphinctering, tongue retraction, larynx height, pitch and larynx tilt interact, is being investigated using simultaneous laryngoscopy and ultrasound (Moisik and Esling 2011a, Esling and Moisik 2011).

These experimental articulatory phonetic findings lay the groundwork for the auditory description of infant speech production. In the past, descriptions of sound qualities emanating from the laryngeal vocal tract have lacked the phonetic precision afforded by the Laryngeal Articulator Model. As a consequence, the labelling of infant speech sounds has been impressionistic rather than auditory phonetic (Stark, Rose, and McLagen 1975, Boysson-Bardies, Sagart, and Durand 1984, Boysson-Bardies and Vihman 1991, Hallé, Boysson-Bardies, and Vihman 1991, McCune, Vihman, Roug-Hellichius, Delery, and Gogate 1996, Oller 2000). Our goal has been to integrate the present findings, based on the Laryngeal Articulator Model, with the findings of these previous studies. The research findings summarized in this paper provide a phonetic account of the components of the laryngeal mechanism, including effects on vowel quality, that can be used to characterize infant vocalizations. It has been proposed that [+cet] is the feature that captures the articulatory generalization of the laryngeal constrictor, including pharyngeal consonants, tonal register, and quality harmony, by creating an independent designation for the action of the lower part of the vocal tract (Czaykowska-Higgins, Moisik, and Esling 2011). With this knowledge, it is possible to ask new questions about the ontogenetic development of the speech modality and to introduce new conjectures about phylogeny.

## 3 Speech Sounds in the First Year of Life

To determine how infants acquire the phonetic modality of speech production, a research project has been established to observe infant vocalizations during the first year of life in four separate language contexts where the phonological inventories of the adult languages contrast. The initial hypothesis is that speech-production processes begin in the laryngeal vocal tract rather than the oral vocal tract. In our model, 'the pharynx is in the larynx.' That is, the pharyngeal articulator is the laryngeal sphincter – the most essential valve complex in the mechanism of the larynx. Therefore, glottal phonation cannot be separated from

pharyngeal activity in the analysis of infant speech. A secondary hypothesis is that manner of articulation is 'learned' at one place of articulation first and then transferred to subsequent places of articulation. The model predicts that stop, approximant, fricative, and trill can all be acquired at the pharyngeal place of articulation before the same manners of oral sounds appear. A tertiary hypothesis is that infants exposed to ambient languages that exploit the laryngeal vocal tract in their phonological inventories will use selected elements of pharyngeal speech production earlier in their babbling behavior than do infants whose ambient languages do not contain laryngeal phonological components. The reasoning is that if speech sounds are acquired first in the pharynx, then it is possible that those articulations will form part of early babbling sequences and earliest phonological acquisition where the phonologies require them.

## 3.1    The InSpA Project: Approach and Method

The four language contexts in the InSpA (Infant Speech Acquisition) Project, where phonological inventories of the ambient languages are designed to contrast, are English in Canada and French in France, which do not have pharyngeal features, and Arabic in Morocco and Bai in Yunnan, China, which do have pharyngeal features. Arabic has two pharyngeal consonants, and Bai has tonal register in which glottal state and laryngeal constrictor setting interact with pitch across the 8 oral syllable categories and 7 nasal syllable categories. Arabic also has emphatic consonants, but these have a different secondary pharyngeal configuration from the primary pharyngeals /ħ/ and /ʕ/ (Hassan and Esling 2011) and will require separate investigation. We are aware of course that English uses laryngeal constrictor settings for paralinguistic purposes and that the /r/ of traditional Parisian French dialect is pharyngeal (and that a pharyngeal reflex also occurs paralinguistically). However, accents of Parisian French have changed over the decades so that it is less likely than before that this feature will be found in the speech of adults, let alone infants in a targeted sampling of this type. The data reported here will cover only results based on the English, Arabic, and Bai samples.

The initial hypothesis tested with our data is that speech-production processes begin in the laryngeal vocal tract rather than the oral vocal tract. Findings in the initial phase of the study are based on data from 19 infants' vocalizations (4 English, 9 Arabic, 6 Bai), classified by trained phoneticians using auditory analysis supplemented by wide-band spectrograms of 3,197 utterances (English: 932; Arabic: 1,011; Bai: 1,254). While not all infants were recorded each month, the data include a recorded session of at least one infant per month from 1 to 12 months. Findings from a later phase of the study are based on approximately 3600 vocalizations, approximately 1200 in each language group, 300 in each of four age groupings: 1-3 months, 4-6 months, 7-9 months, 10-12 months (Benner 2009).

Vocalizations are defined as utterances separated from other utterances by at least two seconds of silence. Digital cameras with integrated microphones were used to record the infants interacting with caregivers in their home environments. Phonatory settings (constricted vs. unconstricted) occurring in babbling sequences towards the end of the first year were also extracted from the data, but all CV(C) utterances, whether monosyllabic, reduplicative, or variegated, were considered babbling and became part of the second-year phase of the Project. This continuing phase of the study aims to pursue the occurrence of pharyngeals developmentally for the particular language groups that have not had wide coverage in the phonological literature – Arabic and Bai. Some findings are mentioned here for the first few months of the second year of life for Arabic.

## 3.2    Exploring Phonatory Capacity: Dynamic Alternations

Infants appear to use their speech production ability to explore laryngeal settings actively and systematically, beginning with phonation, based on physiological predispositions but clearly under the infants' control. In vocalizing, infants are observed to 'play' by manipulating degrees of constriction and increasing the length of utterances while also implementing changes in pitch. Some infants also alternate phonetic parameters dynamically in sequence and then continue to alternate them in increasingly rapid succession over the months. For example, at 5 months, an infant may vary between a phonatory setting with less constriction to a setting with more constriction in a rudimentary pattern of alternation. At 7 months, an infant may alternate phonatory settings with and without constriction; or vocalizations at high pitch without constriction (falsetto) may alternate with high pitched vocalizations with constriction (tight harsh voice). At 8 months, an infant may explore degrees of glottal opening, alternating breathy phonation with modal phonation. At 10 months, constricted events may vary with unconstricted sequences, e.g. harshness with laryngeal closure followed by breathy voice, as unconstricted babbling sequences proliferate. It is the timing-control aspect of these alternations, as well as the nature of the phonetic values being alternated, that suggests to us that these patterns constitute a controlled 'intentional' activity. This patterned kind of alternating play has been more common in infants whose parents encourage the diversity of their vocalization, so to some extent this behavior is conditioned. Nevertheless, it constitutes an intriguing early use of a primal articulatory mechanism, and it deserves to be studied psycholinguistically in more detail.

## 3.3    Speech Begins in the Pharynx

From an anatomical point of view, laryngeal constriction as an available phonetic mechanism can be considered to be predisposed in the infant (Fitch and Giedd

133

1999). In our analysis of the phonetic production data (Bettany 2004, Esling, Benner, Bettany, and Zeroual 2004, Benner, Grenon, and Esling 2007, Grenon, Benner, and Esling 2007), even babies who are surrounded by English produce the laryngeal/pharyngeal sounds that have been identified in the languages described in section 2 that have pharyngeal consonant series, tonal register phonatory contrasts, or vowel harmony secondary coloring. All infants from all four ambient language environments in the study demonstrate the phonetic principle that speech production develops from a laryngeal/pharyngeal basis. The 'first sound' that infants can be said to produce, from a stricture point of view and other than crying with a constricted (retracted) vowel, is epiglottal stop [ʡ], as in Figure 2, since this stricture is a function of the laryngeal constrictor and because this maneuver is the primary airway-protection reflex (Gauffin 1977). Glottal stops [ʔ], which require more careful control than epiglottal stops, allowing slightly more laryngeal openness, only begin to appear in month 2 or 3 in our data. In terms of phonation type, infants begin by using phonatory configurations where laryngeal constriction dominates, e.g. harsh voice, whispery voice, and creaky voice. Gradually, while continuing to employ constriction, infants develop phonation types that do not have constriction, e.g. modal voice, breathy voice, and falsetto. In the earliest months, laryngeally constricted production dominates in all languages observed. Analyses of the initial 3,197 utterances, contrasting only constricted vs. unconstricted utterances across age groups, are significant ($\chi^2(3) =$ 93.34, $p < .001$), indicating that the incidence of laryngeal constriction in infants' vocalizations varies primarily as a function of age, irrespective of linguistic background. In all language groups, earliest vocalization is almost exclusively constricted, that is, pharyngealized. Open-airway phonetic realizations occur only rarely. This is physiologically explainable. Constricted gestures protect the airway, as the infant must be able to close off the airway instantaneously, and the control of pharyngeals is inherently close to the innate reflex response. Producing 'open' (non-constricted) sounds could be viewed as a risk and must be learned gradually. Eventually, all infants progress towards finer control of oral (largely unconstricted) articulations. Nevertheless, there are differences across language environments over the course of the first year in how much constriction is retained and in how many unconstricted types of vocalizations are learned.

**Figure 6:** Proportion of constricted vs. dynamic vs. unconstricted vocalizations produced by Bai infants over the first year of life, in 3-month groupings.
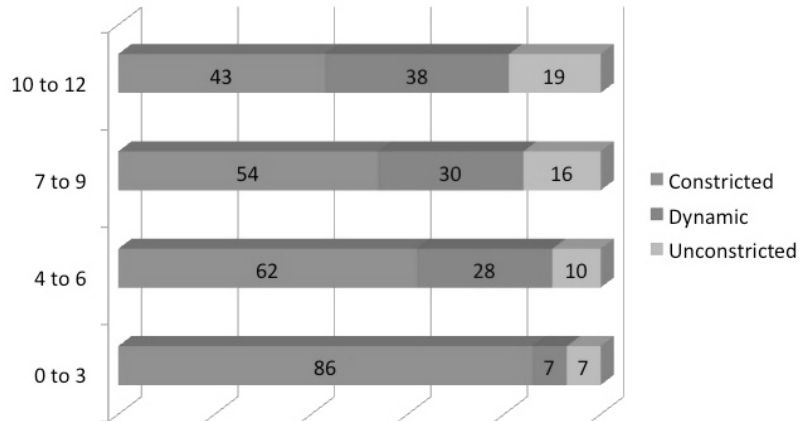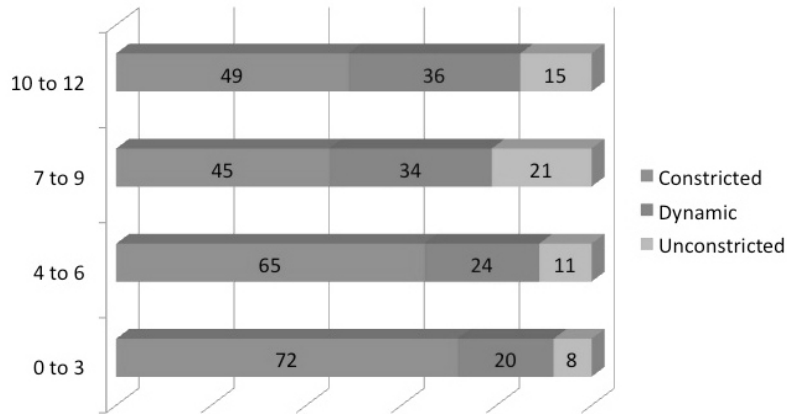


**Figure 7:** Proportion of constricted vs. dynamic vs. unconstricted vocalizations produced by Arabic infants over the first year of life, in 3-month groupings.



A comprehensive analysis contrasting percentages of constricted, dynamic, and unconstricted vocalizations in the larger 3600-utterance sample across age groups in the first year is shown in Figure 6 for the Bai infants. Most vocalizations in the first six months are in the constricted category or vary dynamically between constriction and another sound. Unconstricted vocalizations are rare. Even in the second six months, unconstricted forms constitute less than a fifth of utterances. The same situation is true for Arabic (Figure 7). The move to unconstricted forms does not expand greatly even at the end of the first year.

**Figure 8:** Proportion of constricted vs. dynamic vs. unconstricted
vocalizations produced by English infants over the first year of life,
in 3-month groupings.
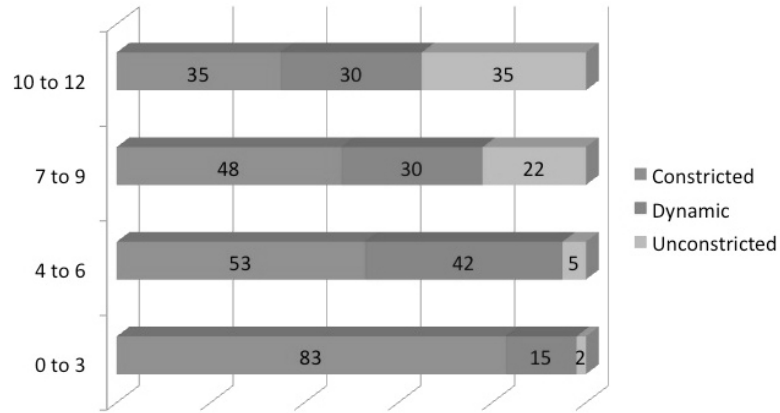


The situation for English infants is the same as for Bai and Arabic infants in the first six months: virtually all sound-producing articulations are in the constricted mode (Figure 8). By the end of their first year, however, there is an increasing tendency for the infants in an ambient English environment to adopt unconstricted forms. This preference for oral sounds may be a product of the lack of pharyngeals (of anything back of velar, in fact) in English phonology. Benner (2009) has demonstrated that English adults/parents show a lack of preference for constricted vocalization in early infant speech, whereas Bai adults/parents show a marked preference for constricted sounds in any infant's vocalizations. This suggests that adult/parent response may play a role influencing which sounds infants are favored to produce (and which sounds lose favor, in the case of an ambivalent or a negative adult response). This could be further tested perceptually. In any case, the fact that infants even in an English context begin with pharyngeals and constricted phonation through most of their first year, just as Arabic or Bai infants do, supports the hypothesis that consonantal and vocalic sounds that are laryngeally constricted (through the action of the pharyngeal articulator) are acquired first, and lead phonetic acquisition, before unconstricted phonation types and oral strictures are able to be acquired.

**3.4    Expanding Place of Articulation: Pharyngeal Priming**

An even more important phonetic process than infants alternating their new phonetic 'discoveries' is what we have called 'pharyngeal priming.' We have observed that infants systematically explore new places of articulation by 'priming' the new sounds with a pharyngeal 'starter,' that is, new oral places of articulation are primed with a preceding sound emanating from the pharyngeal articulator in the larynx. In the first 6 or 7 months, infants produce mainly stop

strictures and fricatives that arise from laryngeal constriction (pharyngeals and glottals). In the second half of the first year, infants begin to 'specialize' in the use of consonants that predominate in their ambient languages. So, in addition to constricted phonation types, infants learn various manners of stricture at the epiglottal/pharyngeal place of articulation. Stop, fricative, approximant, and trill manners are all observed to occur early at the laryngeal constrictor in the pharynx, that is, at the epiglottal/pharyngeal place of articulation (see Figure 1).

Beyond the pharynx, we have found that oral places of articulation do not develop spontaneously as purely oral articulations. Instead, velars, uvulars, labials, alveolars and palatals appear first in combination with a pharyngeal sound. The pharyngeal is typically a continuant. The deeper event primes the higher, oral event. For example, a constriction in the pharynx (harshness, friction, aryepiglottic trilling or just pharyngeal resonance) may precede and prime the oral stricture of a labial [m]; the pharyngeal event acting as a secondary articulation to the labial. Even oral stops may have a distinctly constricted phonation type as an onset or coarticulated background accompaniment (rather than just modal phonation). Fricatives at various places have been observed to be primed with pharyngeal-stricture onset. In months 7–12, infants play by producing sequences to explore all manners at each particular place of articulation (manners which were learned initially in the pharynx). We have ample evidence of infants 'playing with the phonetic instrument', for instance producing uvular trills in long sequences once the uvular place of articulation is discovered, even where uvulars are not found in the ambient language (English).

## 3.5    Constriction in Non-Babbling vs. Babbling

Once babbling begins, babbled utterances tend to be less constricted (more oral) than co-occurring non-babbled vocalizations. In our results thus far, English babbling is the least likely to feature laryngeal constriction, even at an age when the English infants' utterances overall contain a high proportion of constriction. Given that English does not use laryngeal constriction contrastively, this finding was expected. We also found that Arabic babbling consistently included a higher proportion of laryngeally constricted voice quality settings than English babbling.

**Figure 9:** Proportion of constricted vs. dynamic vs. unconstricted non-babbling vocalizations produced by English, Arabic, and Bai infants in the last 3 months of the first year of life.
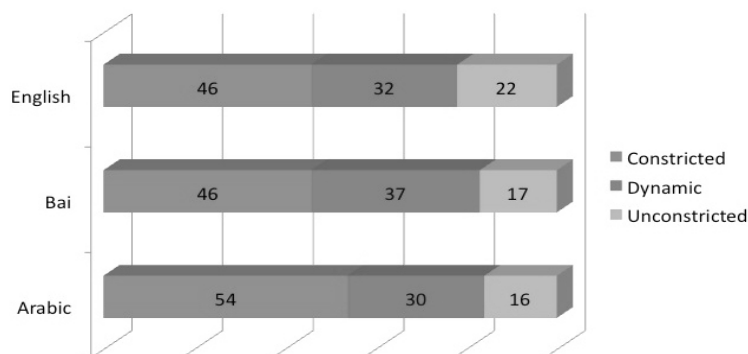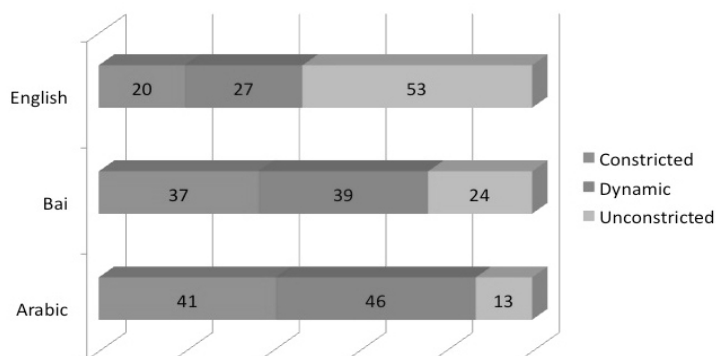


**Figure 10:** Proportion of constricted vs. dynamic vs. unconstricted babbling sequences produced by English, Arabic, and Bai infants in the last 3 months of the first year of life.



The predominance of constricted sounds as laryngeal articulations in infant speech gains further support by observing the comparison between non-babbled vocalizations and early babbling at the end of the first year. Figure 9 shows the percentages of constricted, dynamic, and unconstricted non-babbled vocalizations produced in the three language contexts in months 10-12. Even in the English context, the shift to unconstricted (non-pharyngeal) non-babbled sound types remains around a fifth of the utterances produced. Figure 10 shows the incidence of constricted, dynamic, and unconstricted babbling sequences produced in the three language contexts in months 10-12. For Arabic, the shift to oral sounds is heavily moderated by accompanying (dynamically combined) sequences, and purely oral (unconstricted) sounds are few. For Bai, unconstricted sequences are more numerous, but a large proportion of sounds remain dynamically alternating. For English, there is a clear (53%) preference for oral (unconstricted) sounds by

the time babbling accelerates. Dynamic combinations are fewer, and the incidence of laryngeally constricted forms in babbling sequences is significantly reduced.

This finding is intriguing, since laryngeal constriction in Arabic contrasts mainly at the segmental level. It is possible that the Arabic infants have noticed the parameter 'laryngeal constriction' in their ambient language but have not yet differentiated between segmental and suprasegmental features. Alternatively, it is possible that voice quality settings are more constricted in language contexts where laryngeal constriction is employed at the segmental level. The rate of laryngeal constriction in the babbling of the Bai infants exhibits the most complex pattern by months 10-12. The incidence of constriction in Bai babbling is similar to that of the Arabic infants, while dynamic variation and purely unconstricted babbling are more evenly distributed. It may be that phonetic acquisition of laryngeal contrasts follows a more varied path in Bai than in English or Arabic, given that Bai employs a range of pitch-dependent phonatory settings in its tonal register system, contrasting harsh voice (laryngeally constricted), modal voice, breathy voice, and five pitch patterns (Edmondson, Esling, Lama, Harris, and Li 2001, Esling and Edmondson 2002). Where and when these phonetic contrasts first begin to appear contrastively in meaningful syllables in Bai will require observations of production from the second year of life.

## 4       Discussion

Our first hypothesis – that speech-production processes begin in the laryngeal vocal tract rather than the oral vocal tract – has strong support. The secondary hypothesis, that manner of articulation is 'learned' at the pharyngeal place of articulation first and then transferred to higher oral places of articulation appears to be true, as vocalizations from the laryngeal articulator accompany many other types of sounds. Alternations occur initially with phonation types, exercising states of the larynx, before patterns of oral combining can be performed. Furthermore, it is in the laryngeal region that these alternating patterns learn to be manipulated more and more quickly as control over the articulators grows. Then, sounds that are learned at new, oral places of articulation occur first with clear secondary 'accompaniment' from the original laryngeal/pharyngeal articulator in the performance of 'pharyngeal priming.' The tertiary hypothesis, that infants exposed to ambient languages that exploit the laryngeal/pharyngeal vocal tract in their phonological inventories will use selected laryngeal elements of speech earlier in their babbling behavior than do infants whose ambient languages do not contain laryngeal/pharyngeal phonological components, has only weak support at this stage. By the end of the first year, Arabic and Bai infants have begun to babble using some pharyngeal segments or clearly identifiable laryngeal registers in their CV(C) forms. However, the incidence of these forms is not as great as the proportion of purely unconstricted forms used by the English infants in their

babbling. We have not yet been able to identify, from the data from the end of the first year or from the beginning of the second year, that the Arabic or Bai infants have begun to use pharyngeal segments or clearly identifiable laryngeal registers in meaningful (lexical) contexts. This will require further study with colleagues in Morocco and in China to determine when laryngeal/pharyngeal forms first begin to occur in specific words. It may be that infants whose ambient languages contain laryngeal/pharyngeal phonological components have to 'reacquire' these forms once the acquisitional 'clock' is right. Although the phonetic capacity to perform the pharyngeal/laryngeal forms has already been mastered, and mastered by all infants during the first several months, the actual use of that capacity may need to be activated by other conditions that do not coincide with the beginning of babbling. However, it is also possible that the amount of data we have collected is not yet adequate to give a clear picture of just how many constricted vs. unconstricted forms occur once babbling begins or when first words are beginning to be produced. The amount of speech by 12 months is much greater than the number of vocalizations produced before 10 months, and our methods have been more tooled for phonetic observation than to track phonological development. Hence, it is hypothesized that the emergence of pharyngeals/laryngeals in meaningful contexts will occur during events perhaps very early in the second year of life. Further tracking of months 13-15 and 16-18, particularly for Arabic and Bai, will help to determine whether or not there is a disjuncture in continuity as far as overt articulatory speech production by infants is concerned. In any case, our tracking of the use of constricted forms in the first months of infants' production of speech sounds provides new data to be added to the reformulation of the early stages of the continuity hypothesis (Vihman, Macken, Miller, Simmons, and Miller 1985). Our data collection is continuing into the second year, and analyses with native-language fieldworkers of the production data in the one-word stage in both Arabic and Bai contexts are being designed to address these theoretical issues.

## 5     Conclusions

There are four main points that need to be emphasized in describing the process of the first steps in the acquisition of the speech modality. The first is that we 'learn speech from the inside out;' and we do virtually all of it in the first several months of life. The 'inside' is the 'laryngeal vocal tract' component of the articulatory model, and the 'outside' is the 'oral vocal tract' component of the articulatory model. Laryngeal (including initially pharyngeal and then subsequently glottal) sounds are acquired first; then oral sounds begin to be learned, still employing a laryngeal 'base,' but not necessarily in a predictable order of places of articulation. The 'base' has three aspects: first, the laryngeal base serves as the point at which manner of articulation is acquired; secondly, the laryngeal base serves as the

'pivot point' against which other sounds are practiced in sequences of componential learning; and thirdly, the laryngeal base serves as a motor platform against which new sounds at the oral places of articulation are generated articulatorily.

Secondly, we have shown that manner of articulation can be adequately acquired in the pharynx. This has become clear as a result of our research into various languages with laryngeal phonological distinctions, and as a result of our research on infant speech, where stop, fricative, trill and approximant articulations emerge first as pharyngeals before appearing at later, oral places of articulation. A third finding is that productions of new vocalizations are 'tested' in patterns of alternating sequences that become increasingly supple in their temporal fluidity over the months of practice by the infants. For example, vocalizations with laryngeal constriction are 'tested' in alternation with new vocalizations that do not have laryngeal constriction in patterns of increasingly rapid switching of the articulatory parameters. Finally, we have established that babbling is 'primed' by the innately (i.e. initially reflexively) acquired laryngeal components of the speaking modality, so that the primarily oral sounds that occur typically in babbling emerge not as isolatedly oral articulations at the given place of articulation but rather are 'jump started' by a background, 'secondary' vocalization originating at the laryngeal place of articulation.

With regard to the babbling stage of infant speech, in the second part of the first year, babbling generally prefers new, oral, unconstricted sounds over those sounds that occurred in the first part of the first year and which are characterized primarily by laryngeal constriction. This phenomenon may relate to the split between brain stem neural control and cortical neural control, where brain stem control can be posited to account for the reflexive emergence of the innate use of the laryngeal articulator (responsible for the production of 'pharyngeal' and 'epiglottal' sounds in the taxonomy of the IPA). Later, cortical control is hypothesized to coincide with the shift from phonetic prebabbling practice (as described above) to the primarily oral control exhibited in the babbling stage, beginning around 6 to 8 months. From babbling into the word-development and word-combining stages in the second year of life, it may well be that we 'rediscover' the more basic, innate sounds we learned initially (if they are needed in the target language). Thus, a possible new working hypothesis is that the original, innately acquired constricted sounds will re-emerge 'from the outside in' in languages where they have to emerge eventually in the developmental phonology. The fundamental phonetic principle to be articulated here, on the basis of our research into early articulatory function and control, is that 'motor-phonetic awareness' is acquired very early – arguably just as early as perceptual awareness is known to be acquired, although on a different scale (Kuhl, Stevens, Hayashi, Deguchi, Kiritani, and Iverson 2006). Furthermore, speech production abilities

accompany perceptual abilities in a complex array of motor performance that begins as early as the first few weeks of life.

These results provide evidence for universal and language-specific patterns in the use of laryngeal constriction in the first year of life. In the first months of life, all infants produce universally constricted voice quality settings. Over the course of the year, as infants systematically explore their evolving phonetic capacities, unconstricted settings make up an increasing proportion of their vocal repertoires. Within this general pattern, the distribution of laryngeal constriction in the infants' utterances may vary according to laryngeal/pharyngeal features exploited in the infants' ambient language. A language like English, lacking such features, may develop oral (non-constricted) sounds earlier in babbling. It is estimated that, within the first year of life, infants have learned the major parametric phonetic oppositions of laryngeal constriction vs. openness, glottal closing vs. glottal openness, glottal shortening vs. glottal stretching, larynx raising vs. larynx lowering, and the intermediate degrees required for contrasting manners of articulation, including differing requirements for airflow. The performance data suggest that one-year-old infants can exert paired control over these parameters, to the extent that they already have the articulatory tools at their disposal to produce the phonetically contrastive elements that occur in phonologies that exploit the pharynx, as in Semitic or Tibeto-Burman languages. Notwithstanding, of the range of prelinguistic vocalizations produced by infants, babbling behavior is the most likely to reflect emergent phonological properties and to exhibit cross-linguistic differences.

These findings also raise the question of whether the control of the laryngeal articulator, as a primary and reflexive instrument, may have deeper-seated physiological associations than the oral articulatory gymnastics that are most predominant in babbling. This is a question that will have to be addressed in cognitive science. The various voice qualities that are employed in distinguishing regional and social dialect often have a distinctive laryngeal component (Laver 1980). Phonation type, often erroneously thought of as merely a glottal-level phenomenon, has a close physiological relationship with pharyngeal adjustments (Esling and Harris 2005). The sociolinguistic variety in long-term voice quality, even just in English, has been shown to be as broad as it is finely discriminated (Esling and Edmondson 2011). It is not well known at what age children adopt the specific voice qualities that are part of the regional and social repertory of the accent they are acquiring, and of course without knowing this it is hard to identify if or when children may deviate from the voice quality setting combinations that adults use around them. With our current knowledge that the range of articulatory possibilities in the lower vocal tract, including phonatory effects, is mastered so early in life, we could hypothesize that the use of these features as coarticulatory accompaniments, even to the level of quasi-permanent voice quality, constitutes an element of phonological acquisition that needs to be assessed and analyzed

before we can be certain of the path that phonological acquisition in fact takes. It is at least by no means certain that phonological acquisition can be tracked adequately by noting only the most obvious oral phonetic attribute of the sounds that an infant or a child is producing without noting their laryngeal/pharyngeal coarticulatory features.

It is worth making some comments about phylogeny. Burling's (2005) account of the evolution of language (and speech) makes many points that are strengthened by our research. The assertion that human speech sounds have conventional meaning rather than just being iconic from an early stage seems to be supported by our account of phonetic ontogeny. What our research adds to the equation is that infants acquire motor control over contrastively useful parcels of speech at a surprisingly early age and in a reflexively rich but visually hidden part of the vocal tract. And any speculation about oscillatory patterns of articulators – a concept drawn from MacNeilage (1998) – needs to take into account that these patterns would have developed in the pharynx first, before they progressed to the jaw or the mouth. This provokes speculation about early hominids. Basically, if speech sounds start ontogenetically in the pharynx, then there is every reason to suspect that speech sounds were prominent phylogenetically in the pharynx. In reflecting on Burling's account, it is important to point out that we are talking about infants in both cases rather than adults. It seems likely that speech representations did not start with an early hominid who had already reached adulthood. It would seem that, at any point in time, the path of developing the speech forms that would be used for spoken language was always a question of an infant interacting with adults, predominantly a mother. In our methodology, it has become very clear that adults become intensely aware of the human sound-producing capability when they have infants who are generating the basic elements of phonetic motor production during the first several months of life. The elements are familiar to the adults, but the infant is the driving force; that is, the sounds are created by each infant, in a logical progression of how sounds can be produced in the pharynx, rather than being 'taught' to the infant. Another way of putting this is that we all learn phonetics 'experimentally' (cf. Catford 2001). Early hominid infants, once they had the required cognitive criteria for language development that Burling enumerates, could be expected to have generated sounds similar to those pharyngeal sounds that every infant generates today, which could have served for linguistic meaning; and the mechanism for drawing the phonetic and the semantic processes together would have been precisely because of the infant-mother interaction. Burling's observation that 'it is the parent, not the child, who is the imitator' (2005:110) is given support from our observations of each infant's remarkable control and early mastery of our innate sound-producing instrument.

To summarize, we have identified the pharynx as the origin of earliest speech vocalization and the site of the earliest acquisition of manners of articulation. We

have found evidence of dynamic alternation and strong evidence of pharyngeal priming as active strategies in the speech acquisition process. Data from the Infant Speech Acquisition (InSpA) Project typify instances of 'phonetic play' that demonstrate how infants systematically acquire basic control over the speech mechanism and the arrays of place and manner of articulation during their first year of life. The strategy of pharyngeal priming characterizes how infants progress articulatorily from a laryngeal/pharyngeal base to the formation of oral sounds. We have identified a gradual preference for oral articulations over laryngeal units as babbling progresses, depending on the oral-laryngeal balance in the target phonology. Our initial hypothesis that pharyngeal sounds will proliferate linearly when the ambient language contains those sounds (as in Semitic or in Tibeto-Burman languages) is only weakly supported by the evidence of laryngeals/pharyngeals in late vocalizations in Arabic and in Bai and in the babbling patterns of those languages. We are systematically monitoring the incidence of laryngeals/pharyngeals in Arabic and Bai infants through the second 12 months of life to explore how and when pharyngeal sounds enter production during babbling and in word formation in languages that have pharyngeal (laryngeal constrictor) contrasts.

## Acknowledgements

## References

Benner, Allison. 2009. Production and Perception of Laryngeal Constriction in the Early Vocalizations of Bai and English Infants. Ph.D. dissertation, University of Victoria, Victoria, BC.

Benner, Allison, Izabelle Grenon, and John H. Esling. 2007. Infants' Phonetic Acquisition of Voice Quality Parameters in the First Year of Life. *Proceedings of the 16th International Congress of Phonetic Sciences*, vol. 3, 2073-2076, Saarbrücken.

Bettany, Lisa. 2004. Range Exploration of Pitch and Phonation in the First Six Months of Life. M.A. thesis, University of Victoria, Victoria, BC.

Boysson-Bardies, Bénédicte de, Laurent Sagart, and Catherine Durand. 1984. Discernible Differences in the Babbling of Infants According to Target Language. *Journal of Child Language* 11:1-15.

Boysson-Bardies, Bénédicte de, and Marilyn M. Vihman. 1991. Adaptation to Language: Evidence from Babbling and First Words in Four Languages. *Language* 67:297-319.

Burling, Robbins. 2005. *The Talking Ape: How Language Evolved*. Oxford: Oxford University Press.

Carlson, Barry F., John H. Esling, and Jimmy G. Harris. 2004. A Laryngoscopic Phonetic Study of Nlaka'pamux (Thompson) Salish Glottal Stop, Glottalized Resonants, and Pharyngeals. In D. B. Gerdts and L. Matthewson, eds., *Studies in Salish Linguistics in Honor of M. Dale Kinkade, Occasional Papers in Linguistics No. 17*, 58-71, Missoula: University of Montana Press.

Catford, J. C. 2001. *A Practical Introduction to Phonetics* (2nd edn.). Oxford: Oxford University Press.

Catford, J. C. and John H. Esling. 2006. Articulatory Phonetics. In K. Brown, ed., *Encyclopedia of Language and Linguistics* (2nd edn.), vol. 9, 425-442, Oxford: Elsevier.

Czaykowska-Higgins, Ewa, Scott R. Moisik, and John H. Esling. 2011. Formalizing Lingual-Laryngeal Contrasts: The Epilaryngeal Articulator. Paper presented at 50 Years of Linguistics at MIT, A Scientific Reunion, December 9-11, 2011.

Edmondson, Jerold A., and John H. Esling. 2006. The Valves of the Throat and their Functioning in Tone, Vocal Register, and Stress. *Phonology* 23:157-191.

Edmondson, Jerold A., John H. Esling, Ziwo Lama, Jimmy G. Harris, and Li Shaoni. 2001. The Aryepiglottic Folds and Voice Quality in the Yi and Bai Languages: Laryngoscopic Case Studies. *Mon-Khmer Studies* 31:83-100. [*Minzu Yuwen* 2000(6):47-53 (in Chinese)].

Edmondson, Jerold A., John H. Esling, Jimmy G. Harris, and Huang Tung-chiou. 2005. A Laryngoscopic Study of Glottal and Epiglottal/Pharyngeal Stop and

Continuant Articulations in Amis: An Austronesian Language of Taiwan. *Language and Linguistics* 6:381-396.

Edmondson, Jerold A., Cécile M. Padayodi, Zeki Majeed Hassan, and John H. Esling. 2007. The Laryngeal Articulator: Source and Resonator. *Proceedings of the 16th International Congress of Phonetic Sciences*, vol. 3, 2065-2068, Saarbrücken.

Esling, John H. 1996. Pharyngeal Consonants and the Aryepiglottic Sphincter. *Journal of the International Phonetic Association* 26:65-88.

Esling, John H. 1999. The IPA Categories 'Pharyngeal' and 'Epiglottal': Laryngoscopic Observations of Pharyngeal Articulations and Larynx Height. *Language and Speech* 42:349-372.

Esling, John H. 2005. There are No Back Vowels: The Laryngeal Articulator Model. *Canadian Journal of Linguistics* 50:13-44.

Esling, John H. 2006. States of the Glottis. In K. Brown, ed., *Encyclopedia of Language and Linguistics* (2nd edn.), vol. 12, 129-132, Oxford: Elsevier.

Esling, John H. 2010. Phonetic Notation. In W. J. Hardcastle, J. Laver and F. E. Gibbon, eds., *Handbook of Phonetic Sciences* (2nd edn.), 678-702, Oxford: Blackwell.

Esling, John H., Allison Benner, Lisa Bettany, and Chakir Zeroual. 2004. Le contrôle articulatoire phonétique dans le prébabillage. *Actes des XXVes Journées d'Étude sur la Parole*, 205-208, Fès, Maroc: AFCP.

Esling, John H., and Jerold A. Edmondson. 2002. The Laryngeal Sphincter as an Articulator: Tenseness, Tongue Root and Phonation in Yi and Bai. In A. Braun and H. R. Masthoff, eds., *Phonetics and its Applications: Festschrift for Jens-Peter Köster on the Occasion of his 60th Birthday*, 38-51, Stuttgart: Franz Steiner Verlag.

Esling, John H., and Jerold A. Edmondson. 2011. Acoustical Analysis of Voice Quality for Sociophonetic Purposes. In Marianna Di Paolo and Malcah Yaeger-Dror, eds., *Sociophonetics: A Student's Guide*, 131-148, London & New York: Routledge.

Esling, John H., Katherine E. Fraser, and Jimmy G. Harris. 2005. Glottal Stop, Glottalized Resonants, and Pharyngeals: A Reinterpretation with Evidence from

a Laryngoscopic Study of Nuuchahnulth (Nootka). *Journal of Phonetics* 33:383-410.

Esling, John H., and Jimmy G. Harris. 2005. States of the Glottis: An Articulatory Phonetic Model Based on Laryngoscopic Observations. In W. J. Hardcastle and J. M. Beck, eds., *A Figure of Speech: A Festschrift for John Laver*, 347-383, Mahwah, NJ: Lawrence Erlbaum.

Esling, John H, and Scott R. Moisik. 2011. Multimodal Observation and Measurement of Larynx Height and State during Pharyngeal Sounds. *Proceedings of the 17th International Congress of Phonetic Sciences*, 643-646. Hong Kong.

Fitch, W. Tecumseh, and Jay Giedd. 1999. Morphology and Development of the Human Vocal Tract: A Study using Magnetic Resonance Imaging. *Journal of the Acoustical Society of America* 106:1511-1522.

Gauffin, Jan. 1977. Mechanisms of Larynx Tube Constriction. *Phonetica* 34:307-309.

Grenon, Izabelle, Allison Benner, and John H. Esling. 2007. Language-Specific Phonetic Production Patterns in the First Year of Life. *Proceedings of the 16th International Congress of Phonetic Sciences*, vol. 3, 1561-1564, Saarbrücken.

Hallé, Pierre A., Bénédicte de Boysson-Bardies, and Marilyn M. Vihman. 1991. Beginnings of Prosodic Organization: Intonation and Duration Patterns of Disyllables Produced by Japanese and French Infants. *Language and Speech* 34:299-318.

Hassan, Zeki Majeed, and John H. Esling. 2011. Investigating the Emphatic Feature in Iraqi Arabic: Acoustic and Articulatory Evidence of Coarticulation. In Zeki Majeed Hassan & Barry Heselwood, eds., *Instrumental Studies in Arabic Phonetics,* 217-234, Amsterdam: John Benjamins.

Hassan, Zeki Majeed, John H. Esling, Scott R. Moisik, and Lise Crevier-Buchman. 2011. Aryepiglottic Trilled Variants of /ʕ, ħ/ in Iraqi Arabic. *Proceedings of the 17th International Congress of Phonetic Sciences*, 831-834, Hong Kong.

Heselwood, Barry. 2007. The 'Tight Approximant' Variant of the Arabic 'ayn. *Journal of the International Phonetic Association* 37:1-32.

Kuhl, Patricia K., Erica Stevens, Akiko Hayashi, Toshisada Deguchi, Shigeru Kiritani, and Paul Iverson. 2006. Infants Show a Facilitation Effect for Native Language Phonetic Perception between 6 and 12 Months. *Developmental Science* 9(2):F13-F21.

Laver, John. 1980. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

MacNeilage, Peter F. 1998. The Frame/Content Theory of Evolution of Speech Production. *Behavioral and Brain Sciences* 21:499-511.

McCune, Lorraine, Marilyn M. Vihman, Liselotte Roug-Hellichius, Diane Delery, and Lakshmi Gogate. 1996. Grunt Communication in Human Infants (Homo sapiens). *Journal of Comparative Psychology* 110:27-37.

Moisik, Scott R. 2008. A Three-Dimensional Model of the Larynx and the Laryngeal Constrictor Mechanism: Visually Synthesizing Pharyngeal and Epiglottal Articulations Observed in Laryngoscopy. M.A. thesis, University of Victoria, Victoria, BC.

Moisik, Scott R., and John H. Esling. 2011a. Evaluating the Vowel Space Effects of Larynx Height using Laryngeal Ultrasound. *Canadian Acoustics* 39:180-181.

Moisik, Scott R., and John H. Esling. 2011b. The 'Whole Larynx' Approach to Laryngeal Features. *Proceedings of the 17th International Congress of Phonetic Sciences*, 1406-1409, Hong Kong.

Moisik, Scott R., John H. Esling, and Lise Crevier-Buchman. 2010. A High-Speed Laryngoscopic Investigation of Aryepiglottic Trilling. *Journal of the Acoustical Society of America* 127(3):1548-1559.

Oller, D. Kimbrough. 2000. *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum.

Padayodi, Cécile M. 2008. Kabiye. *Journal of the International Phonetic Association* 38:215-221.

Stark, Rachel E., Susan N. Rose, and Margaret McLagen. 1975. Features of Infant Sounds: The First Eight Weeks of Life. *Journal of Child Language* 2:205-221.

Tiede, Mark K. 1996. An MRI-Based Study of Pharyngeal Volume Contrasts in Akan and English. *Journal of Phonetics* 24:399-421.

Vihman, Marilyn M., Marlys A. Macken, Ruth Miller, Hazel Simmons, and Jim Miller. 1985. From Babbling to Speech: A Re-Assessment of the Continuity Issue. *Language* 61:397-445.

John H. Esling
Department of Linguistics
University of Victoria
PO Box 1700
Victoria BC  V8W 2Y2
Canada

esling@uvic.ca